Markov Chains and Associated Algorithms: A Compilation of Some Proofs and Solutions
Daniel Grossman
Computer Science and Mathematics 1999

Abstract

The theory of randomized algorithms constitutes a large field of study in computer science, as such solutions often demonstrate average time complexities far superior to those of non-randomized alternatives. Under Professor Ravi Kannan, I have become interested in the mathematical theory of one type of randomized process, random walks on graphs, and the synonymous theory of Markov chains ("discrete time stochastic process[es] defined over a [finite] set of states S in terms of a matrix P of transition probabilities," where the matrix P has one row and column for each state S [Lovász]).

In my first of two semesters of work on the subject, my first goal was to closely read and understand a small number of important treatments of the topic. These included "Markov Chains and Random Walks" from the text *Randomized Algorithms* (Motwani and Raghaven, 1995), and *Markov Chains and Polynomial Time Algorithms* (Kannan, 1994). With the background and advanced knowledge supplied therein, I proceeded with the main goal of the project: to closely read, understand, and supply the omitted proofs and solutions in Laszlo Lovász's work "Random Walks on Graphs: A Survey" (1993).

In my first semester thesis, I began with a summary of some introductory theoretical material on the subject matter. I continued by providing proofs for several assertions which Lovasz leaves "to the reader" to justify, and also solutions to some problems that he speculates would be "interesting for the reader" to solve. The areas covered included access time on paths and circuits, cover time, bounds on some of these main parameters, and the connection between Markov chains and the eigenvalues of a graph's Markov matrix. My study of Lovász's work and the field in general continued in the second semester of my thesis with a look at commute time, graph symmetry, and more eigenalgebra, and several applications of that and the previous semester's work.

1 Introduction to Markov Chains

Lovász begins his work (on page 353 of the Bolyai Society's anniversary edition: *Combinatorics, Paul Erdos is Eighty*) with a short description of what a random walk on a graph is, and how it pertains to a Markov chain.

For a graph G, consider the act of picking a starting point and then selecting a neighbor of this point to move to at random. Then pick another new node at random and move there. The random path that we traverse in such a fashion (specifically, the sequence of nodes that we *walk*) is known as a *random walk* on the graph.

In Lovász's work (and therefore, here as well), attention is mostly restricted to the case of undirected graphs without weighted edges. As a result, every random walk we perform is precisely a finite, time-reversible Markov chain. While Markov chains often have varying weights on their edges, we will see that our edges will be walked with uniform probability from every node; that is, the probability of taking any edge out of a given vertex will be the same as that for any other edge.

1.1 Definitions

Before discussing the results of random walks, let us first define several useful variables and traditional symbols.

Our graph $G = (V, E)$ is assumed to be connected and simple, with $|V| = N$ nodes and $|E| = M$ edges. Lovász does briefly touch on problems concerning graphs without these properties, but we will not do so here.

As noted above, the random steps we take in our walks will be uniformly distributed at each node. We express this as: given a *starting node* v_0 and the node (v_t) we have reached after our t^{th} step (or *on* our t^{th} step, depending on whether one thinks of the destination nodes or the edges as the point of incrementing the step count), the probability of moving to each neighbor of v_t is $1/d(v_t)$, where $d(i)$ is the degree of the node i. Note that there is no differentiation between *in-degree* and *out-degree* because our graph is undirected. Also note that the probability that $v_{t+1} = v_t$ is 0, because our graph is simple.

Now, the initial node v_0 may be fixed in some walks, or may be drawn from an initial probability distribution P_0 over the set V. Given a distribution from which our first node is drawn, v_t is not specified; instead we refer to the distribution of v_t, denoted P_t:

$$P_t(i) = \text{Prob}(v_t = i)$$

We use $M = (p_{ij})_{i,j \in V}$ to denote the matrix of *transition probabilities* from node i to node j in our walk. Clearly, from our definitions so far, p_{ij} has value $1/d(i)$ if the edge ij exists, and has value 0 otherwise.

Another method of realizing this *transition matrix* is apparent. First, we define A_G to be the adjacency matrix of G, and we let D denote the diagonal matrix with entry $(D)_{ii} = 1/d(i)$. Then, $M = DA_G$; each node j adjacent to node i has transition probability equal to the degree of i, and those not neighboring i have 0 transition probability. If G is d-regular, then there is no need for the degree matrix — all the degrees are the same — and so we can write $M = (1/d)A_G$. But we will not deal with such matrices.

Now, then, the *rule* of the walk is readily available in vector-matrix form:

$$P_{t+1} = M^T P_t,$$

with the probability distributions at each time-step expressed as vectors. Following this rule back through time gives us an expression for the probability distribution at any time t based solely on the initial distribution and the transition probabilities:

$$P_t = (M^T)^t P_0$$

An important consequence of this simple equation is that the probability (p^t_{ij}) of reaching node j, starting from node i, in time t (i.e. in t "steps"), is given directly by entry ij of the matrix M^t.

As Lovász points out, the successive probability distributions $P_0, P_1, \ldots$ are not the same; rather, they vary with time. However, if P_1 *is* the same as P_0 (and thus $P_t = P_0$ for all $t \geq 0$), then we are engaged in the *stationary walk.* This does not mean that we remain on the starting node, but rather that, given an initial point distribution, our walk will take us to each of the points with the same probability as there was initially of our being there.

With the knowledge that such a stationary walk is possible, we define $\pi(v)$ to be the *stationary distribution*, where

$$\pi(v) = \frac{d(v)}{2m} \tag{1}$$

How is this distribution known to be stationary? Clearly, if the probability of arriving at each node happens to be the degree of that node (i.e. the number of edges running into the node) divided by the total number of edges, then we will be at each node with the same probability at each time step t. We can also see that this distribution is stationary by the operation of the transition matrix on π:

$$\left[\frac{d(i)}{2m}\right]^T * DA_G = \pi$$

As we will show, the stationary distribution is important because non-bipartite graphs demonstrate convergence of the random walk distributions to π as $t \to \infty$.

1.2 Main Parameters

The following are some of the important values that characterize random walks; only some of them are utilized in this paper.

The *access time* or *hitting time* $H(i, j)$ in a random walk is the expected number of steps needed to reach node j from node i.

The bidirectional sum $H(i, j) + H(j, i)$ is the *commute time* in a random walk — that is, the expected number of steps to go from node i through node j and back to i again.

The *cover time* $C(i)$ is the expected number of steps necessary to reach every node in the graph, where i is the starting node. The cover time can also be discussed in terms of a starting distribution.

Finally, the *mixing rate* μ measures how many steps are required before the random walk converges to a limiting distribution, if one exists.

2 Discussion of Some Parameters

Lovász proceeds to illustrate how to solve for several of these parameters under certain conditions, and also how to apply them. We look at a handful of his examples, especially those that require some filling in.

2.1 Access Time

Lovász continues on page 358 with an example (#1) demonstrating the notion of access time. We are to calculate the access time $H(i, k)$ for two arbitrary nodes, $0 \leq i < k < n$, on a path of nodes $0, 1, \ldots, n-1$. First, Lovász observes that the access time $H(k-1, k)$ is equivalent to the expected return time of a random walk on a path with $k+1$ nodes, starting at an end node, minus one.

Why is this the case? Return time refers to the number of steps taken after leaving a node before returning to that node. If we begin on node $k-1$ in the above path, then "returning to" node k (i.e. the $(k+1)^{\underline{st}}$ node) will take one fewer steps than had we started on node k, since we have already eliminated one step from our journey by beginning on node $k-1$.

Now, the return time of the endnode of a path of $k+1$ nodes can be shown to be $2k$. If we are engaged in a random walk, then the expected return time to any node v is $\pi(v)^{-1} = \frac{2m}{d(v)}$. Since in the case of an endnode $d(v) = 1$, the return time equals $2m$ — in this case, we have k edges, so the return time is $2k$. So, $H(k-1, k) = 2k - 1$. Thus, we have the following recurrence:

$$H(i, k) = H(i, k-1) + (2k-1), \tag{2}$$

because in order to reach node k from node i, we must first reach node $k-1$, which requires $H(i, k-1)$ steps.

Lovász sums this recurrence to obtain

$$H(i, k) = k^2 - i^2 \tag{3}$$

and notes that $H(0, k) = k^2$ (Brownian Motion). Also, we see that the cover time of the path on n nodes, starting from 0, is equivalent to $H(0, n-1) = (n-1)^2 - (0)^2 = (n-1)^2$.

Lovász continues by asserting that the access time between two nodes at a distance k in a *circuit* of length n is $k(n-k)$. We can demonstrate that this is, in fact, the correct solution. First, we show that the recurrence for obtaining this result has only one solution, using linear independence. Note that since the first step from node i is to $i+1$ with probability 1/2 and to $i-1$ with probability 1/2, we have:

$$H(i, i+k) = \frac{1}{2}(H(i-1, i+k) + 1) + \frac{1}{2}(H(i+1, i+k) + 1) \tag{4}$$

$$H(i, i+k) = \frac{1}{2}H(i-1, i+k) + \frac{1}{2}H(i+1, i+k) + 1 \tag{5}$$

Clearly $H(i, i+k)$ depends only on k (and not on i). So, assigning the variable H_k to the access time between two points at distance k, we thus have:

$$\begin{aligned} H_k &= \frac{1}{2}H_{k+1} + \frac{1}{2}H_{k-1} + 1, \quad \text{for} 1 \leq k \leq \frac{n}{2} && (6) \\ 0 &= \frac{1}{2}H_{k-1} - H_k + \frac{1}{2}H_{k+1} + 1 && (7) \end{aligned}$$

Thus, to show that there is only one solution for H_k, it suffices to show that the $n-2$ equations $0 = \frac{1}{2}H_{k-1} - H_k + \frac{1}{2}H_{k+1} + 1, \quad 0 < k < n-1$ are linearly independent; that is, a linear combination of these $n-2$ equations equals the $\mathbf{0}$ vector if and only if each equation is multiplied by 0. We begin with $k = 1, k = 2, k = 3, \ldots$ (i.e. the top of the system):

$$\begin{aligned} \frac{1}{2}H_0 - H_1 + \frac{1}{2}H_2 &= -1 && (8) \\ \frac{1}{2}H_1 - H_2 + \frac{1}{2}H_3 &= -1 && (9) \\ \frac{1}{2}H_2 - H_3 + \frac{1}{2}H_4 &= -1 && (10) \\ &\vdots \end{aligned}$$

Note that since the H_i's may vary, the only coefficient (in a linear combination) that can be applied to equation (8) to guarantee a $\mathbf{0}$ sum is 0, since H_0 appears only here. Then we see that the only possible means to generate a $\mathbf{0}$ vector element in the H_1 slot is by multiplying (9) by 0, as that is the only remaining equation containing H_1. We can continue this process for all the $n-2$ equations that we generated from the recurrence, showing that they all must have the 0 coefficient in order to generate a 0-sum linear combination. Thus, they are linearly independent.

At this point we have shown that there can be only one solution to the access time recurrence relation. Thus, if Lovász's solution is accurate, then it is the only solution for $H(i, i+k)$ on a circuit of length n. We verify this now, using the recurrence in (4):

$$\begin{aligned} &\frac{1}{2}[(k+1)(n-(k+1))] + \frac{1}{2}[(k-1)(n-(k-1))] + 1 \\ &= \frac{1}{2}[kn - k^2 - k + n - k - 1] + \frac{1}{2}[kn - k^2 + k - n + k - 1] + 1 \\ &= kn - k^2 \\ &= k(n-k), \quad \text{as required} && (11) \end{aligned}$$

2.2 Cover Time from an Internal Node

Lovász next leaves as an exercise to the reader the following problem — determine the cover time of the path on n nodes $(0, 1, \ldots, n-1)$ when starting from an internal node i.

To solve this problem, we note the Markov recurrence that describes the situation, with $f(i)$ representing the cover time from some node i:

$$f(i) = \frac{1}{2}f(i-1) + \frac{1}{2}f(i+1) + 1, \quad \text{for } 1 \leq i \leq n-2 \qquad (12)$$

This makes sense because, for any internal node, covering the entire path requires making a first step either to the left or to the right. But once we have taken this step, we see that if we continue to the left endnode, it will still remain to traverse the entire path back to the right endnode, and vice versa. In other words, there are two ways to cover a path starting at an internal node, depending solely on which endnode is encountered first. But the cover time from either endnode is $(n-1)^2$, as proved earlier. So, we have two base cases for the above recurrence:

$$f(0) = f(n-1) = (n-1)^2 \tag{13}$$

Let us try to obtain a closed form expression for this function f by analyzing the successive difference in its value at different points along the path. First, we rearrange the terms of the recurrence relation to obtain

$$\begin{aligned} \frac{1}{2}[f(i) - f(i-1)] + \frac{1}{2}[f(i) - f(i+1)] &= 1 & (14) \\ f(i) - f(i-1) &= 2 + [f(i+1) - f(i)] & (15) \end{aligned}$$

Starting at one end of the path, say node $n-1$, let us denote the value of the successive difference between that node and the one to the left of it as x. That is,

$$f(n-1) - f(n-2) = x \tag{16}$$

Then, it is clear from equation (15) that $f(n-2) - f(n-3) = x + 2$, and so on down the path. But, due to symmetry, we know that when we reach the other endpoint, 0, the sum of all the successive differences must cancel, such that the cover time from 0 is once again *equivalent* to the cover time from $n-1$. In other words, adding the x terms and the increasing multiples of 2 from each fragment along the way, the sum of the successive differences is

$$\begin{aligned} (n-1)x + \sum_{i=1}^{n-1} 2(i-1) &= (n-1)x + \left[\frac{2(n-2)}{2}(n-1)\right] \\ &= (n-1)x + (n-2)(n-1) \end{aligned}$$

Setting this sum to 0, then, we obtain $x = -(n-2)$. Recalling what x was substituted for, we have

$$\begin{aligned} f(n-1) - f(n-2) &= -(n-2) \\ f(n-1) + (n-2) &= f(n-2) \\ (n-1)^2 + (n-2) &= f(n-2) \end{aligned}$$

Similarly, $f(n-3) = (n-1)^2 + 2(n-2)$, etc., and we arrive at a closed form expression for f:

$$f(n-i) = (n-1)^2 + (n-2)(n-i), \quad \text{for } 1 \le i \le n-2$$

To express f as a function of the starting node i, we substitute $(n-i)$ and reach the final form for our cover time function:

$$f(i) = (n-1)^2 + (n-2)(n-i-1), \quad \text{for } 1 \le i \le n-2$$

2.3 Cover Time on a Circuit

On the bottom of page 358, Lovász next determines the cover time of the circuit on nodes $0, 1, ..., n-1$. The point here is that this covering is equivalent in elapsed time to reaching n nodes of a long *path*, starting at its midpoint and including this midpoint.

Why should this be so? Note that if we traverse this long path such that every new node we visit is to the left of our original node, then after $n-1$ new nodes, we have traversed a distance equivalent to that of the entire circuit we are modeling. Similarly with the right half of the long path. So, allowing for crossover between the two halves, we see that the entire circuit is covered in the same time as $n-1$ new nodes are covered on the path.

Clearly, then, in order to cover n nodes, we must first cover $n-1$ nodes. (This type of recurrence argument is, as may be evident, common to Markov analysis.) We have:

$$C(n) = C(n-1) \ + \ \text{the time to then reach the last new node} \tag{17}$$

What is the time to reach that final node? Recall that we are trying to find one new node on a *very long path*. Thus, we have the option of moving to the adjacent node or moving to the next new node at the other end of the $(n-1)$-node subpath we have covered so far. Now we can use the reverse reasoning from before, where we unfolded a circuit into a long path; that is, we fold this long path back into a circuit. It is clear that walking to the next new node, either the adjacent one or the one "far away" at the other end of the path, is identical to accessing a node in an n-length circuit from one adjacent to it. We can either take one step, or walk all the way around the circle.

Now, then, we already have the solution for this problem: the access time between two nodes at distance 1 apart on a circuit of length n is $1(n-1) = n-1$, from equation (11). So, Lovász's recurrence from (17) is completed:

$$C(n) = C(n-1) + (n-1); \quad \text{where } C(1) = 0, \tag{18}$$

and expanding this equation to the sum of the integers from 1 to $n-1$, we achieve the same expression for the cover time as he:

$$C(n) = \frac{n(n-1)}{2} \tag{19}$$

2.4 Further Exploration of the Main Parameters

Lovász's other basic example, regarding access times for the complete graph on n nodes, K_n, is thoroughly explained in the text; we therefore do not discuss it here. Next, Lovász presents some bounds on the parameters of access time, commute time, and cover time and cites several papers for proof. These bounds are not essential to our current paper, but Lovász brings up several points that we explore. First, he notes the effect of symmetry on access times. Clearly, his Proposition 2.2 is true: *If u and v have the same degree, then the probability that a random walk starting at u visits v before returning to u is equal to* *the probability that a random walk starting at v visits u before returning to v.* Since u and v have the same degree, all probabilistic calculations regarding entering and leaving these nodes remain the same upon permuting u and v.

As Lovász notes, if the degrees of the two nodes are different we must observe that the first probability in Proposition 2.2 is proportional to the probability of visiting v under the stationary distribution (and vice versa for the second probability). That is, as the probability of being at v increases, the probability that our Markov chain will hit it before returning home increases. Thus, we can form a ratio of the probabilities of "intermediate visitation" as $\pi(v)/\pi(u) = d(v)/d(u)$.

But if we are interested in the exact probability that node v is hit before returning to node u, how should we proceed? Lovász's explanation centers around determining the ratio of the return time to u versus the commute time through v and back to u. First, we note that this is an adequate probability measure; since the first time of return, say τ, is necessarily less than or equal to the first time of return after visiting v, say σ, the ratio of their expected values is less than one. But does it really measure the probability we're looking for? We want to find $q = P\{\tau = \sigma\}$. A helpful hint Lovász provides is that if $\tau < \sigma$, then after the first τ steps, in order to attain σ we must still commute through node v, which we have not yet hit. In other words, we can condition on whether or not $\sigma = \tau$:

$$\begin{aligned} E(\sigma - \tau) &= E(\sigma - \tau | \tau = \sigma) P\{\tau = \sigma\} + E(\sigma - \tau | \tau < \sigma) P\{\tau < \sigma\} \\ &= (0 * q) + E(\sigma) * (1 - q) \end{aligned} \tag{20}$$

because the expected value of $\sigma - \tau$ (given that we have already returned to u but yet not hit v) is precisely the commute time through $v - \sigma$. As such, we have that

$$\begin{aligned} E(\sigma) - qE(\sigma) &= E(\sigma - \tau) \\ qE(\sigma) &= E(\sigma) - E(\sigma) + E(\tau) \\ q &= \frac{E(\tau)}{E(\sigma)} \end{aligned} \tag{21}$$

and by definition

$$\begin{aligned} q &= \frac{2m}{d(u)\kappa(u,v)} \\ &= \frac{1}{\kappa(u,v)\pi(u)} \end{aligned} \tag{22}$$

Let us skip Theorem 2.4 in the text for now. We will return to it and a corollary (and also to an introductory unproven remark by Lovász from the first section) with knowledge gleaned from Section 3 below. We instead conclude this section with Lovász's next topic, determining cover time from access time for a random walk. His first theorem here states upper and lower bounds on the cover time for a random walk, given the maximum and minimum access times (respectively) between any two nodes. These results are rigorously proved by Matthews elsewhere, and we instead focus on proving a weaker upper bound than the theorem states — i.e. $2log_2n$ times the maximum access time, rather than $1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n} = log_2n$ times the maximum access time.

First, Lovász has us note his Lemma 2.8: *Let b be the expected number of steps before a random walk visits more than half of the nodes, and let h be the maximum access time between any two nodes. Then $b \leq 2h$.* In fact, we can show $b < 2h$ strictly, as follows. First, assume as the author does that $n = 2k + 1$ (thus odd). Denote α_v as the first time node v is visited. Then the time β to reach more than half of the nodes is the middle, or $(k+1)$st largest of the α_v. Now then, even if we visited a new node on every step, minimizing the $\sum_v \alpha_v$, we see via arithmetic progression that this sum would be greater than or equal to $(1 + max(\alpha_v))\frac{n}{2} \geq k + k * max(\alpha_v) \geq k * 2\beta$, or as Lovász gives, $(k+1)\beta$. So, as defined above, $b = E(\beta) \leq \frac{1}{k+1} \sum_v E(\alpha_v)$. Since the maximum access time is h, $b \leq \frac{nh}{k+1} = \frac{n}{k+1}h < 2h$.

Again towards proving the $2log_2n$ bound we were looking for, note that in time $2h$ we have seen more than half of the nodes. As if restarting our walk, we will again cover more than half the nodes in time $2h$. Since half of these are expected to be ones we have already seen, we are in fact covering only $\frac{1}{2} * \frac{1}{2}$ new nodes (or half the remaining nodes) in this second $2h$ amount of time. Inductively, the time to cover the entire graph is thus $2h$ times log_2n time intervals, or $2log_2n$ times the maximum access time, as we desired to show.

3 Eigenvalues and Eigenvectors

In moving on to Lovász's third section, let us recall the meaning of the random walk transition matrix M. It was formed by the product of the diagonal matrix D — with each D_{ii} set to the reciprocal of the degree of node i — and A_G — the adjacency matrix of the graph G upon which we are performing our random walks.

Because the probability of moving from node i to node j in one step is contained in element (i, j) of the transition matrix M (by definition), it is clear that repeated multiplication (say, t times) by this *random walk operator* will let us read off the probability of moving from i to j in that many steps, straight from the resulting matrix. That is, the probability of p_{ij}^t of starting at i and moving to j in t steps is now element (i, j) of the matrix M^t (times whatever the initial distribution was).

3.1 Largest Eigenvalue; The Frobenius-Perron Theorem

Now, what properties does the original transition matrix M possess? First, we can determine the value of its largest eigenvalue, which will be an important parameter for considering the effects of the matrix M^t in the limit as $t \to \infty$. Recall that the definition of eigenvalues states that $Mx = \lambda x$ for all eigenvalues λ and vectors x in the n-dimensional space. Given a vector x, consider the index i such that $|x_i| \geq |x_j|$ for all $j \neq i$. Then we have

$$\sum_{j=1}^{n} M_{ij}x_j = \lambda x_i \tag{23}$$

because the expression on the left is precisely a multiplication of the $i^{\underline{th}}$ row of the matrix with the vector, generating the $i^{\underline{th}}$ row (element) of the product vector. This gives us

$$\begin{aligned} |\lambda||x_i| &= |\sum_j M_{ij}x_j| &(24)\\ &\leq \sum_j M_{ij}|x_j| &(25)\\ &\leq |x_i|\sum_j M_{ij} &(26) \end{aligned}$$

because $|x_i|$ was the maximum element. And since the sum of any row in the Markov transition matrix must be 1 (meaning that at every step, we must leave whatever node we are currently on and head for another one), we see that

$$|\lambda||x_i| \leq |x_i| * 1 \tag{27}$$

So, $|\lambda| \leq 1$. This result can be expanded upon via the Frobenius-Perron theorem of linear algebra, which (for our purposes) assures that there will be some maximal eigenvalue with an entirely positive-coordinate corresponding eigenvector. Taking advantage of this property (which holds for irreducible, non-negative matrices only), we can find the largest eigenvalue by multiplying M by the entirely positive-coordinate vector of all 1's:

$$M * [1 \cdots 1] \quad \text{or} \quad M * \mathbf{1} \tag{28}$$

which simply sums the row entries of M. As stated before, the rows of the matrix must sum to 1, giving

$$\begin{aligned} M * \mathbf{1} &= \lambda * \mathbf{1} &(29)\\ \mathbf{1} &= \lambda * \mathbf{1} &(30)\\ \lambda &= 1 &(31) \end{aligned}$$

Thus, the maximal eigenvalue for the transition matrix M is 1. Note that we have also found a right eigenvector for this λ, namely $\mathbf{1}$.

3.2 The Dominant Eigenvector and the Stationary Distribution

Next, recall that the stationary distribution π for a random walk over the nodes of the graph is achieved as the limiting value of the the distributions as t increases. Since M has largest eigenvalue = 1, our intuition would be that the corresponding left eigenvecotr would be precisely that stationary distribution, π.

Why? Since the maximal eigenvalue reflects (as noted before) what happens in the limit, as the matrix is applied repeatedly to some initial distribution vector, we expect the stationary distribution to be generated. And so it is: multiplying the matrix M by the vector π simply generates π again, for by definition, once the stationary distribution is reached, it cannot be left through any number of applications of M. Thus, we have

$$\pi^T * M = \pi^T \tag{32}$$

$$\pi^T * M = \pi^T * 1 \tag{33}$$

And since 1 is an eigenvalue, π is clearly the corresponding left eigenvector.

3.3 Convergence of p_{ij}^t to the Stationary Distribution

After establishing these fundamental properties regarding the transition matrix, Lovász continues in an effort to prove the convergence of $p_{ij}^t \to \pi(j)$.

3.3.1 Matrix Symmetry; Application of the Spectral Theorem

First, in order to make useful assumptions about the matrices with which we are dealing, Lovász goes about generating a symmetric version of the transition matrix (since it cannot be guaranteed symmetric as is unless the graph G is known to be regular). Recalling the diagonal matrix of vertex degrees, D, and the adjacency matrix of the graph, A_G, we consider the matrix contruction $N = D^{1/2} A_G D^{1/2}$. Because A_G **is** a symmetric matrix (since for every node i adjacent to node j, node j is adjacent to i) and D is a diagonal matrix, N must be symmetric as well. With $M = DA_G$, we can substitute into our construction:

$$N = D^{-1/2} M D^{1/2} \tag{34}$$

Due to the symmetry we now have, we can express N in its *spectral form* (as per the Spectral Theorem of Matrix Algebra):

$$N = \sum_{k=1}^{n} \lambda_k v_k v_k^T \tag{35}$$

where $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ are the eigenvalues of N and $v_1, \ldots, v_n$ are the corresponding eigenvectors of *unit length*. (I do not offer a proof of the Spectral Theorem here. Please consult an algebra text such as Algebra, Michael Artin, pp. 253.) Constructing a vector w where $w_i = \sqrt{d(i)}$, we have an eigenvector (*not* of unit length) of N with eigenvalue = 1. This is clear through multiplication:

$$\begin{aligned} &[\cdots \sqrt{d(i)} \cdots] D^{-1/2} M D^{1/2} \\ &= \mathbf{1}^T M D^{1/2} \\ &= \mathbf{1}^T D^{1/2} \\ &= [\cdots \sqrt{d(i)} \cdots] \end{aligned} \tag{36}$$

Lovász next appeals to Frobenius-Perron to demonstrate that $\lambda_1 = 1 > \lambda_2 \geq \cdots \geq \lambda_n \geq -1$. To unitize the w vector arrived at above, and thus obtain v_1, we divide w by its norm, $\sqrt{d(1)^2 + \cdots + d(n)^2} = \sqrt{2m}$, thus giving $v_1 = 1/\sqrt{2m}$. Lovász notes that each element of this unit eigenvector (corresponding to λ_1) is $\sqrt{\frac{d(i)}{2m}}$, or $\sqrt{\pi(i)}$.

3.3.2 Aside: Connection Between λ_n and Graph Bipartiteness

Lovász further asserts that if G is non-bipartite, then λ_n is strictly greater than -1, but does not prove this non-trivial fact. How can we arrive at the same conclusion? First, we know that matrices M and N have the same eigenvalues because they are related by the similarity transformation $N = D^{-1/2}MD^{1/2}$; we therefore discuss M's eigenvalues, aware that they are also N's.

Let us then prove that if M has $\lambda_n = -1$, then G *is* bipartite, arriving at Lovász's statement by the contrapositive. For w an eigenvector of M, then, $Mw = -w$. Let us one again pick an index i such that $|w_i|$ is maximum in w. Clearly,

$$\sum_j M_{ij} = 1, \tag{37}$$

as previously noted with regard to the eigenvector of all 1's. Then, by definition, we have

$$\begin{aligned} |(Mw)_i| &= |\sum_j M_{ij} w_j| & (38) \\ |w_i| &\leq \sum_j M_{ij} |w_j| & (39) \end{aligned}$$

Dividing and distributing through the sum, we can trap the $|w_j|$ values between two inequalities:

$$1 \leq \sum_j M_{ij} \left| \frac{w_j}{w_i} \right| \tag{40}$$

Since $|w_i|$ is maximal in w, all of the $\left|\frac{w_j}{w_i}\right|$ terms are less than or equal to 1. But we know that without these terms, the row would sum to 1, as in equation (37) above. Thus, if any of the fraction terms were less than 1, the sum would *not* be greater than or equal to 1. So, we have that by definition $\left|\frac{w_j}{w_i}\right| \leq 1$ for all j, but by equation (40) $\left|\frac{w_j}{w_i}\right| \geq 1$ for all j. The result is that all of the elements' aboslute values must be equivalent: $|w_i| = |w_j| \ \forall i, j$.

How does this help show that the graph must be bipartite? We have assumed from the beginning of this work that the graphs on which we are walking are connected. So, we can examine, for some node i, all nodes j adjacent to i by locating the non-zero entries of row i in M. Now, we can express our prior eigen-equation that reflected the -1 eigenvalue, $Mw = -w$ as

$$\sum_j M_{ij} w_j = -w_i \tag{41}$$

But we have already shown that $|w_j| = |w_i|$ for all j. This means that if for some j, w_j has the same sign as w_i, the total on the left hand side of (41) will be of smaller magnitude than $|w_i|$. In other words, all the w_j's must have opposite sign from w_j or else their "weighted" sum (weighted, that is, byt the entries of M_{ij}) will not amount to the negative magnitude of w_i. If any single w_j is the opposite sign from the rest of the w_j's, it will detract from the sum, and the equality above will not hold.

We can thus define a binary property over the graph (namely, the sign of each node's entry in the w eigenvector) such that every node is adjacent to nodes with the opposite value for the property. Grouping together those nodes with positive entries and those with negative entries gives us precisely a bipartite graph. Thus, if the transition matrix, and then also N, has eigenvalue $\lambda_n = -1$, the graph must be bipartite, showing Lovász's assertion to be true.

3.3.3 Convergence

We continue towards the goal of demonstrating p_{ij}^t convergence by obtaining the matrix containing that value at the intersection of the $i^{\underline{th}}$ row and $j^{\underline{th}}$ column, M^t. First, we see that

$$\begin{aligned} N^t &= D^{-1/2} M D^{1/2} D^{-1/2} M D^{1/2} \ \cdots \text{ t groups } \cdots \ D^{-1/2} M D^{1/2} \\ N^t &= D^{-1/2} M^t D^{1/2} \\ M^t &= D^{1/2} N^t D^{-1/2} \end{aligned}$$

Using the Spectral Theorem once again, we re-express M^t in its spectral form:

$$M^t = \sum_{k=1}^{n} \lambda_k^t D^{1/2} v_k v_k^T D^{-1/2} \tag{42}$$

The case where $k = 1$ is special, as we already know the values of λ_1, v_1, and v_1^T. We then have:

$$\begin{aligned} \lambda_1^t D^{1/2} v_1 v_1^T D^{-1/2} &= 1 * \left[\ddots \frac{1}{\sqrt{\pi(i)}} \ddots \right] \left[\sqrt{\pi(i)\pi(j)} \right] \left[\ddots \sqrt{\pi(j)} \ddots \right] \\ &= [\cdots \pi(j) \cdots] \end{aligned} \tag{43}$$

And so,

$$p_{ij}^t = \pi(j) + \sum_{k=2}^{n} \lambda_k^t D^{1/2} v_k v_k^T D^{-1/2} \tag{44}$$

As before, we see that if G is not bipartite, then $|\lambda_i| < 1$ for $i = 2, \ldots, n$, and so the sum approaches 0 as $t \to \infty$. Thus, we have that

$$p_{ij}^t \to \pi(j) \quad \text{as } t \to \infty \tag{45}$$

As stated before, the stationary distribution π is the left eigenvector for the maximum eigenvalue of 1, and is thus the expected limiting value of repeated applications of M to a starting distribution. We now have further mathematical proof for this conjecture.

3.4 Matrix Spectrum and the Access Time Matrix

On page 387, Lovász begins a deeper look into the connection between matrix spectra and the properties of the random walks they govern. In terms of the matrices involved, he seeks to derive a formula for access times — what he calls a "spectral formula".

First, we define a matrix H over the vertices of the graph G, where H_{ij} = the access time from node i to node j, i.e. $H(i,j)$. Lovász suggests that we proceed by defining a typical Markov recurrence over the set $\Gamma(i)$, the neighbors of node i. For $i \neq j$ and $\{v_1, \ldots, v_n\}$ the members of $\Gamma(i)$, we have:

$$H(i,j) = 1 + \frac{1}{d(i)}H(v_1,j) + \frac{1}{d(i)}H(v_2,j) + \cdots + \frac{1}{d(i)}H(v_n,j) \tag{46}$$

since the first random step takes us from i to any one of the nodes in $\Gamma(i)$ with equal probability and we must then access j from that node. In summation form, we have:

$$H(i,j) = 1 + \frac{1}{d(i)} \sum_{v \in \Gamma(i)} H(v,j) \tag{47}$$

Lovász insightfully expresses this recurrence in a matrix notation, as follows:

$$F = J + MH - H, \quad \text{where J is the matrix of all 1s} \tag{48}$$

This grouping makes sense, as we naturally multiply only the non-zero elements in M (i.e. those that define an adjacency) by the elements of H. This leaves us with a diagonal matrix F, since $J + MH - H$ cancels everything but the $H(i,i)$ terms on the main diagonal. How do we deal with this F?

Multiplying by the left eigenvector of the matrix M, π, we obtain

$$\begin{aligned} F^T\pi &= J\pi + (MH)^T\pi - H^T\pi \\ &= J\pi + H^T\left(M^T - I\right)\pi \\ &= J\pi + H^T\left(M^T - I^T\right)\pi \\ &= J\pi + H^T\left(M - I\right)^T\pi \end{aligned} \tag{49}$$

Since π is the stationary distribution, $M^T\pi = \pi$; also, $I^T\pi = \pi$, giving us

$$F^T\pi = J\pi = \mathbf{1} \tag{50}$$

since the sum of the elements of the stationary distribution, as with all distributions over the vertices of G, is 1.

We can therefore specify F precisely, since we know it consists of the diagonal entries $\frac{1}{\pi(i)}$ such that (50) holds. Thus, given $\pi(i) = \frac{d(i)}{2m}$,

$$F = 2mD \tag{51}$$

And substituting for F, $J + MH - H = J - (I - M)H$, we arrive at

$$(I - M)H = J - 2mD \tag{52}$$

Given this "matrix equation" (as Lovász puts it), we would like to solve for our access time matrix, H. However, this is not immediately possible, because $(I - M)$ matrix. That is, as shown below, there are infinitely many solutions to the equation. Lovász remarks that for any matrix X that solves the equation, every other matrix of the form $X + \mathbf{1}a^T$ also satisfies it, for any vector a. How does he arrive at this result?

Assuming that the equation does have more than one solution, we set up a system between two of them:

$$(I - M)X = J - 2mD \tag{53}$$

$$(I - M)X = J - 2mD \tag{54}$$

Subtracting, we obtain $(I - M)(X - Y) = 0$. Let $G = X - Y$. We would like to say that if $(I - M)G = 0$, then $G = \mathbf{1}a^T$ for some vector a. Distributing and separating terms, we have

$$G = MG \tag{55}$$

which is of the form of an eigenvalue equation. In fact, (55) means precisely that each column of G is an eigenvector of M with eigenvalue $= 1$. But we have already examined the properties of the matrix M, and we know that there exists a unique eigenvector for $\lambda = 1$, namely the vector of all 1s. By definition, then, each of these eigenvector columns of G is simply a scalar multiple of $\mathbf{1}$. Expressed in *tensor product* notation, $G = \mathbf{1}a^T$ for some vector a with unspecified scalars as its elements.

Now, we must solve

$$(I - M)Y = J - 2mD, \quad \text{where } Y = X + \mathbf{1}a^T \tag{56}$$

and a base condition is that $Y(i, i) = 0$. That is only to say that the access time between node i and itself is 0. Now, clearly we can bring the matrix Y to a form where the main diagonal is all 0s by subtracting constant vectors from any column, since all matrices $X + \mathbf{1}a^T$ satisfy the equation. So, if we subtract the scalar value in Y_{ii} from column i for all i, we will have have a matrix Y that solves for H and obeys the main diagonal restriction.

Now, then, Lovász suggests that we look at the matrix M^*, which is defined as the limit of M^t as $t \to \infty$. It follows from several earlier arguments that $M^*_{ij} = \pi(j)$. Then, with

$$\begin{aligned}
I &= D^{1/2}\left(\sum_{k=1}^{n} e_k e_k^T\right) D^{-1/2} \\
M &= \sum_{k=1}^{n} D^{1/2}\lambda_k v_k v_k^T D^{-1/2} \\
M^* &= \sum_{k=1}^{n} D^{1/2}\lambda_k^* v_k v_k^T D^{-1/2} = D^{1/2} v_1 v_1^T D^{-1/2}
\end{aligned}$$

$I - M + M*$ leaves us with an invertible expression:

$$\left(\sum_{k=1}^{n} \lambda_k v_k e_k^T\right)\left(\sum_{k=1}^{n} \frac{1}{\lambda_k} e_k v_k^T\right), \tag{57}$$

and substitution shows $Y = (I - M + M^*)^{-1}(J - 2mD)$ to satisfy equation (56).

Setting $H = (I - M + M^*)(J - 2mD)$ and diagonalizing the matrix M into a spectral form as before leads (after some tedious calculation) to Lovász's final formula for the access time:

$$H(s,t) = 2m \sum_{k=2}^{n} \frac{1}{1-\lambda_k}\left(\frac{v_{kt}^2}{d(t)} - \frac{v_{ks}v_{kt}}{\sqrt{d(s)d(t)}}\right) \tag{58}$$

Well, that took some effort to determine, but it gives us an immediate result for the commute time between nodes s and t as follows:

$$\begin{aligned}
\kappa(s,t) &= H(s,t) + H(t,s) \\
&= 2m \sum_{k=2}^{n} \frac{1}{1-\lambda_k}\left(\frac{v_{kt}^2}{d(t)} - \frac{v_{ks}v_{kt}}{\sqrt{d(s)d(t)}}\right) + 2m \sum_{k=2}^{n} \frac{1}{1-\lambda_k}\left(\frac{v_{ks}^2}{d(t)} - \frac{v_{kt}v_{ks}}{\sqrt{d(t)d(s)}}\right) \\
&= 2m \sum_{k=2}^{n} \left(\frac{v_{kt}^2}{d(t)} - \frac{2v_{ks}v_{kt}}{\sqrt{d(s)d(t)}} + \frac{v_{ks}^2}{d(s)}\right) \\
&= 2m \sum_{k=2}^{n} \frac{1}{1-\lambda_k}\left(\frac{v_{kt}}{\sqrt{d(t)}} - \frac{v_{ks}}{\sqrt{d(s)}}\right)^2
\end{aligned} \tag{59}$$

as Lovász also attains.

As we have seen, $-1 \leq \lambda_k \leq 1$, or $1 - \lambda_k \leq 2$. Thus, $\frac{1}{2} \leq \frac{1}{1-\lambda_k}$. Furthermore, since in our sum above we are dealing with $k \geq 2$, $\lambda_2 \geq \lambda_k$ by definition, and thus $1 - \lambda_2 \leq 1 - \lambda_k$, giving $\frac{1}{1-\lambda_k} \leq \frac{1}{1-\lambda_2}$. Thus, we have lower and upper bounds for $\frac{1}{1-\lambda_k}$.

Given this range, we can compute a range for the commute time function as well. First, for the upper bound:

$$\begin{aligned}
\kappa(s,t) &= 2m \sum_{k=2}^{n} \frac{1}{1-\lambda_k}\left(\frac{v_{kt}}{\sqrt{d(t)}} - \frac{v_{ks}}{\sqrt{d(s)}}\right)^2 \\
&\leq \frac{2m}{1-\lambda_2} \sum_{k=1}^{n} \left(\frac{v_{kt}}{\sqrt{d(t)}} - \frac{v_{ks}}{\sqrt{d(s)}}\right)^2
\end{aligned}$$

because adding the first eigenvector only increases the upper bound, and does not cause division by zero since the eigenvalue term has been maximized and removed from the sum. Then, due to the orthogonality of the matrix and the consequence that the difference between two orthogonal vectors is equal to the sum of their squares, we have

$$\leq \frac{2m}{1-\lambda_2} \sum_{k=1}^{n} \left(\frac{v_{kt}^2}{d(t)} + \frac{v_{ks}^2}{d(s)}\right)$$

$$\leq \frac{2m}{1-\lambda_2}\left(\frac{1}{d(t)}+\frac{1}{d(s)}\right) \tag{60}$$

For the lower bound on the commute time we note (as we could have above but weren't yet motivated to do so!) that the summand for $k = 1$ is zero and so does not incorrectly increase the lower bound when we add it to the sum to take advantage of the orthogonal vector property:

$$\begin{aligned}
\kappa(s,t) &\geq \frac{2m}{2}\sum_{k=1}^{2}\left(\frac{v_{kt}^2}{d(t)}+\frac{v_{ks}^2}{d(s)}\right) \\
&\geq m\left(\frac{1}{d(s)}+\frac{1}{d(t)}\right) \qquad (61)
\end{aligned}$$

Let us examine these bounds, then. For a regular graph the commute time is bounded from below by n. How do we see this? Since regularity gives $1/d(s) = 1/d(t)$,

$$\kappa(s,t) \geq m\left(\frac{2}{d(i)}\right) = \frac{2m}{d(i)}$$

and since $\sum_V d(i) = 2m$ with each $d(i) = \frac{2m}{n}$ the same,

$$\kappa(s,t) \geq \frac{2m}{\frac{2m}{n}} = n$$

For the important class of *expander graphs*, which are regular graphs with the property that $\frac{1}{1-\lambda_2} = O(1)$ (i.e. a constant not depending on n), we see that

$$\begin{aligned}
\kappa(s,t) &\leq 2mc\left(\frac{2}{d(i)}\right) \\
&\leq 2c\left(\frac{2m}{d(i)}\right) \\
&\leq 2cn
\end{aligned}$$

That is, the upper bound on the commute time is $O(n)$. And since expander graphs are regular graphs, the lower bound is also $O(n)$, making the commute time $\Theta(n)$.

3.5 More with Eigenvalues and Eigenvectors

As promised earlier in this paper (and as Lovász recommends at this point) we will now return to some unproven results from the second section and prove them using our newfound eigenvalue/vector techniques. First, we look at Theorem 2.4 on page 362. We try to prove that *for any three nodes u, v, and w, $H(u,v) + H(v,w) + H(w,u) = H(u,w) + H(w,v) + H(v,u)$*. This is known as the *weak symmetry property of access times*. Let us rearrange the terms and prove:

$$(H(u,v) - H(v,u)) + (H(v,w) - H(w,v)) + (H(w,u) - H(u,w)) = 0 \tag{62}$$

The first two terms can be expanded as follows:

$$2m \sum_{k=2}^{n} \frac{1}{1-\lambda_k} \left(\frac{v_{kv}^2}{d(v)} - \frac{v_{ku} v_{kv}}{\sqrt{d(u)d(v)}} \right) - 2m \sum_{k=2}^{n} \frac{1}{1-\lambda_k} \left(\frac{v_{ku}^2}{d(u)} - \frac{v_{kv} v_{ku}}{\sqrt{d(v)d(u)}} \right) \tag{63}$$

Note that the "cross-product" terms cancel straightforwardly and will do so as well for the other two groupings in equation (62). Then, by inspection we see that the three vector square terms will each be added once and subtracted once upon writing the equation out in full, thus leaving us with

$$(62) \; = 2m \sum_{k=2}^{n} \frac{1}{1-\lambda_k} * 0 = 0$$

proving Theorem 2.4.

Before continuing with Lovász's next suggested proof, let us continue with his Corollary 2.5, stating that *the nodes of any graph can be ordered so that if* u *precedes* v *then* $H(u,v) \leq H(v,u)$. *Such an ordering can be obtained by fixing any node* t *and ordering the nodes according to the value of* $H(u,t) - H(t,u)$. Following Lovász's proof, first let us say that u "precedes" v. That is to say that

$$\begin{aligned} H(u,t) - H(t,u) &\leq H(v,t) - H(t,v) \\ H(u,t) + H(t,v) &\leq H(v,t) + H(t,u) \end{aligned} \tag{64}$$

Given Theorem 2.4, though, we can substitute for the left hand side of (64), such that

$$\begin{aligned} H(u,v) + H(v,t) + H(t,u) - H(v,u) &\leq H(v,t) + H(t,u) \\ H(u,v) &\leq H(v,u) \end{aligned} \tag{65}$$

as required.

Now we will return our attention to one of Lovász's first formulas, (2.1) on page 357, in which he expresses the access time between two nodes as a function of the commute time between the two and the sum of the commute times between every node in the graph and the two endpoints in question. More clearly, we are presented with

$$H(i,j) = \frac{1}{2} \left(\kappa(i,j) + \sum_{u} \pi(u)[\kappa(u,j) - \kappa(u,i)] \right) \tag{66}$$

Two methods are suggested for proving this fact, one of them using the electrical resistance formulae in the text (which we do not cover here), and the other — eigenvalues. We will prove it with Theorem 3.1 and Corollary 3.2 regarding access and commute times: $H(i,j) =$

$$\begin{aligned} &= \frac{1}{2}\kappa(i,j) + \frac{1}{2} \sum_{u} \pi(u)[\kappa(u,j) - \kappa(u,i)] \\ &= \frac{2m}{2} \sum_{k=2}^{n} \frac{1}{1-\lambda_k} \left(\frac{v_{kj}}{\sqrt{d(j)}} - \frac{v_{ki}}{\sqrt{d(i)}} \right)^2 & (67) \\ &+ \frac{2m}{2} \sum_{u} \pi(u) \left[\sum_{k=2}^{n} \frac{1}{1-\lambda_k} \left(\frac{v_{kj}}{\sqrt{d(j)}} - \frac{v_{ku}}{\sqrt{d(u)}} \right)^2 - \sum_{k=2}^{n} \frac{1}{1-\lambda_k} \left(\frac{v_{ki}}{\sqrt{d(i)}} - \frac{v_{ku}}{\sqrt{d(u)}} \right)^2 \right] & (68) \end{aligned}$$

Let us expand the terms in this sum separately. First, (67)

$$= m \sum_{k=2}^{n} \frac{1}{1-\lambda_k} \left(\frac{v_{kj}^2}{d(j)} - \frac{2v_{ki}v_{kj}}{\sqrt{d(i)d(j)}} + \frac{v_{ki}^2}{d(i)} \right) \tag{69}$$

and second, (68)

$$\begin{aligned}
&= m \sum_u \pi(u) \left\{ \sum_{k=2}^{n} \frac{1}{1-\lambda_k} \left[\left(\frac{v_{kj}^2}{d(j)} - \frac{2v_{kj}v_{ku}}{\sqrt{d(j)d(u)}} + \frac{v_{ku}^2}{d(u)} \right) - \left(\frac{v_{ki}^2}{d(i)} - \frac{2v_{ki}v_{ku}}{\sqrt{d(i)d(u)}} + \frac{v_{ku}^2}{d(u)} \right) \right] \right\} \\
&= m \sum_u \pi(u) \sum_{k=2}^{n} \frac{1}{1-\lambda_k} \left(\frac{v_{kj}^2}{d(j)} - \frac{v_{ki}^2}{d(i)} - \frac{2v_{kj}v_{ku}}{\sqrt{d(j)d(u)}} + \frac{2v_{ki}v_{ku}}{\sqrt{d(i)d(u)}} \right) \\
&= m \sum_u \frac{d(u)}{2m} \sum_{k=2}^{n} \frac{1}{1-\lambda_k} \left(\frac{2v_{ki}v_{ku}}{\sqrt{d(i)d(u)}} - \frac{2v_{kj}v_{ku}}{\sqrt{d(j)d(u)}} \right) + m \sum_{k=2}^{n} \frac{1}{1-\lambda_k} \left(\frac{v_{kj}^2}{d(j)} - \frac{v_{ki}^2}{d(i)} \right) \qquad (70)
\end{aligned}$$

We can cancel the m terms and distribute the $\frac{1}{2}$ through the first term of (70). Note that we grouped (70) the way we did in order to straightforwardly add (69) back into the equation, cancel the v_{ki}^2 terms, add the v_{kj}^2 terms together, and obtain

$$= \sum_u \sum_{k=2}^{n} \frac{\sqrt{d(u)}}{1-\lambda_k} \left(\frac{v_{ki}v_{ku}}{\sqrt{d(i)}} - \frac{v_{kj}v_{ku}}{\sqrt{d(j)}} \right) + m \sum_{k=2}^{n} \frac{1}{1-\lambda_k} \left(\frac{2v_{kj}^2}{d(j)} - \frac{2v_{kj}v_{ki}}{\sqrt{d(i)d(j)}} \right) \tag{71}$$

Finally, in (71) the first sum is composed of the difference of two fractions, each of which is the inner product of the first eigenvector with the kth eigenvector. Since the eigenvectors form an orthonormal basis, these inner products are all zero and drop out of the equation. Thus we are left with precisely Theorem 3.1's result:

$$H(i,j) = 2m \sum_{k=2}^{n} \frac{1}{1-\lambda_k} \left(\frac{v_{kj}^2}{d(j)} - \frac{v_{ki}v_{kj}}{\sqrt{d(i)d(j)}} \right) \tag{72}$$

and Lovász's (originally Tetali's) access time formula is proven.

As another corollary to 3.1, the text next looks at the averge access time for reaching all nodes t. Since the stationary distribution is the uniform distribution, we can compute an average simply by summing the products of $\pi(t)H(s,t)$ over all t. Lovász's work here on page 369 is more or less straightforward, but a couple eigenvector tricks we have used previously can be redemonstrated here. We have

$$\begin{aligned}
\sum_t \pi(t)H(s,t) &= \sum_t \frac{d(t)}{2m} * 2m \sum_{k=2}^{n} \frac{1}{1-\lambda_k} \left(\frac{v_{kt}^2}{d(t)} - \frac{v_{ks}v_{kt}}{\sqrt{d(s)d(t)}} \right) \\
&= \sum_t \sum_{k=2}^{n} \frac{1}{1-\lambda_k} \left(v_{kt}^2 - v_{ks}v_{kt}\sqrt{\frac{d(t)}{d(s)}} \right) \\
&= \sum_{k=2}^{n} \frac{1}{1-\lambda_k} \left(\sum_t v_{kt}^2 - v_{ks}\sqrt{\frac{1}{d(s)}} \sum_t v_{kt}\sqrt{d(t)} \right)
\end{aligned}$$

Then here again we use the fact that the eigenvectors are of unit length (being an orthonormal basis) and that for $k \geq 2$ (which is precisely the range of our sum), all the v_k are orthogonal to v_1 and therefore orthogonal to each other. Thus, the average access time works out to

$$\sum_t \pi(t) H(s,t) = \sum_{k=2}^{n} \frac{1}{1-\lambda_k} \tag{73}$$

independently of the start node s used for calculation. As Lovász notes earlier and we can plainly see now, the frequency of $1/(1-\lambda_k)$ terms suggest that we would benefit from finding good bounds on the so-called *spectral gap*, that is, $1 - \lambda_2 = \lambda_1 - \lambda_2$. The text does delve into calculations regarding this parameter in the section on electrical resistance, but we do not cover that here.

Let us conclude with the text's final points regarding the transition matrix and its eigenvalues and eigenvectors. In the middle of page 370 we have another application of Theorem 3.1, this time explaining the intuitive fact about random walks (and, by extension, Brownian motion) that "more distant targets are more difficult to reach". We seek to compute the average access time to a fixed node t over all starting nodes s. We proceed in a similar fashion as we did in proving (73):

$$\begin{aligned}
\sum_s \pi(s) H(s,t) &= \sum_s \frac{d(s)}{2m} * 2m \sum_{k=2}^{n} \frac{1}{1-\lambda_k} \left(\frac{v_{kt}^2}{d(t)} - \frac{v_{ks} v_{kt}}{\sqrt{d(s)d(t)}} \right) \\
&= \sum_s \sum_{k=2}^{n} \frac{1}{1-\lambda_k} \left(v_{kt}^2 \frac{d(s)}{d(t)} - v_{kt} v_{ks} \sqrt{\frac{d(s)}{d(t)}} \right) \\
&= \sum_{k=2}^{n} \frac{1}{1-\lambda_k} \left(\frac{v_{kt}^2}{d(t)} \sum_s d(s) - \frac{v_{kt}}{\sqrt{d(t)}} \sum_s v_{ks} \sqrt{d(s)} \right)
\end{aligned}$$

Once again using the orthogonality of the eigenvectors, and that the sum of the degrees of all the nodes is $2m$, we arrive at

$$\begin{aligned}
\sum_s \pi(s) H(s,t) &= \sum_{k=2}^{n} \frac{1}{1-\lambda_k} \left(\frac{v_{kt}^2}{d(t)} * 2m - 0 \right) \\
&= \frac{2m}{d(t)} \sum_{k=2}^{n} \frac{1}{1-\lambda_k} v_{kt}^2 \qquad (74)
\end{aligned}$$

Now what can we say about equation (74)? Lovász notes the similarity between our familiar $\frac{1}{1-\lambda_k}$ term and the terms of arithmetic and harmonic series, and cites the following *inequality between arithmetic and harmonic means*:

$$\frac{\sum_i \frac{1}{x_i} w_i}{\sum_i w_i} \geq \frac{\sum_i w_i}{\sum_i x_i w_i}$$

Viewing the v_{kt}^2 terms as the "weights" w_i in the mean inequality above, we can write

$$\frac{\sum_{k=2}^{n} \frac{1}{1-\lambda_k} v_{kt}^2}{\sum_{k=2}^{n} v_{kt}^2} \geq \frac{\sum_{k=2}^{n} v_{kt}^2}{\sum_{k=2}^{n} (1-\lambda_k) v_{kt}^2} \tag{75}$$

The total weights can be simplified by incorporating adding and subtracting the $k = 1$ element of the sum:

$$\sum_{k=2}^{n} v_{kt}^2 = \sum_{k=1}^{n} v_{kt}^2 - \pi(t) = 1 - \pi(t)$$

We can also simplify the arithmetic sum term, as we know that $\lambda_1 = 1$:

$$\begin{aligned} \sum_{k=2}^{n}(1-\lambda_k)v_{kt}^2 &= \sum_{k=1}^{n}(1-\lambda_k)v_{kt}^2 \\ &= \sum_{k=1}^{n} v_{kt}^2 - \lambda_k v_{kt}^2 \\ &= 1 - \sum_{k=1}^{n} \lambda_k v_{kt}^2 \end{aligned}$$

But then this sum term is simply the (t,t) element of the spectral form of the matrix $N = D^{1/2}AD^{1/2}$, which has all zeroes on the main diagonal. Thus,

$$\sum_{k=2}^{n}(1-\lambda_k)v_{kt}^2 = 1 - (N)_{t,t} = 1$$

Concluding, we place values in (75)

$$\begin{aligned} \frac{\sum_{k=2}^{n} \frac{1}{1-\lambda_k} v_{kt}^2}{1-\pi(t)} &\geq \frac{1-\pi(t)}{1} \\ \sum_{k=2}^{n} \frac{1}{1-\lambda_k} v_{kt}^2 &\geq [1-\pi(t)]^2 \end{aligned}$$

and substitute this last into equation (74) to obtain

$$\begin{aligned} \sum_{s} \pi(s)H(s,t) &\geq \frac{d(m)}{2t}[1-\pi(t)]^2 \\ &\geq \frac{1}{\pi(t)}[1-\pi(t)]^2 \end{aligned} \tag{76}$$

Since for some nodes s close to t the access time will in fact be very small, and yet the *average* access time we just computed may be very large, we see that certain nodes s that are far from t must have very large access times indeed.

4 Conclusion

In this treatment, we have covered a great deal of Lovász's survey of random walks on graphs, and yet we have covered only a fraction of the many applications of Markov chain theory to other stochastic processes, such as Martingales and Brownian motion. Indeed, Markov chain theory is widely used in physical, chemical, and even financial modeling, wherein certain assumptions make the use of random walks appropriate to a model of the process taking place. Often such theoretical applications will stay well away from matrices and instead focus solely on conditioning or Bayesian arguments from pure probability theory. Nonetheless, as was one of our major points in this work, we have seen that matrices and eigenalgebra come in rather handy in determining many theoretical and practical results. Lovász's "survey" of random walks has been quite a thorough introduction and, through our elaboration on it, has allowed us to gain a deeper understanding of the underlying Markov chain, probability, and matrix theory.

References

[1] Motwani, Rajeev, and Prabhakar Raghaven. *Randomized Algorithms*. Cambridge University Press, Cambridge, 1995.

[2] R. Kannan. "Markov Chains and Polynomial Time Algorithms". *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*. pp 656-671. IEEE, Los Alamitos, 1994.

[3] L. Lovász. "Random Walks on Graphs: A Survey". *Combinatorics, Paul Erdos is Eighty, Volume II.* pp 353-397. Bolyai Society for Mathematical Studies, Budapest, 1996.